Land use land cover analysis with pixel-based classification approach

Haslina Hashim, Zulkiflee Abd Latif, Nor Aizam Adnan
Faculty of Architecture, Planning and Surveying, UniversitiTeknologi MARA, Malaysia

Article Info	ABSTRACT
	Rapid development in certain urban area will affect its natural features. Therefore, it is important to identify and determine the changes occur for further analysis and future development planning. This process will influence several factors such as area development, environmental issues and human social activities. The selection of remote sensing data and method will derive the accurate land use land cover maps. This research study accessed the classification accuracy of different classifier approach for land use land cover classification in urban area. The objective of this paper is to compare the accuracy of the classification for each technique used. The study was conducted in a highly urbanized area in Kuala Lumpur, Malaysia. The dataset used for this study is the multi temporal LANDSAT satellite imageries for the year of 2001, 2006, 2011 and 2016. The pre-processing and analysis of the dataset has been done using software ENVI 5.3. Five land use classes (Urban/built up area, Forest, Agriculture, Water Body and fallow land) were identify for classification process. The classification approach for this study is the supervised classification with two algorithms namely Maximum Likelihood (ML) and Support Vector Machine (SVM). The overall accuracy and kappa statistic of the classification indicate that support vector machine algorithm was more accurate than maximum likelihood algorithm for five different time intervals. Therefore, this classification approach is acceptable and highly recommended for mapping the changes of land cover.

Corresponding Author:

Haslina Hashim,
Applied Remote Sensing & Geospatial Research Group,
Faculty of Architecture, Planning and Surveying,
Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia.
Email: haslinah8118@gmail.com

1. INTRODUCTION

Environmental changes mostly occurred in an area that have rapid development and high-density population. Surface cover on the ground such as vegetation area, water elements and development area represent the land cover of an area [1]. Human activities definitely will change the land covers. This activity will give huge impact towards sustainable resources, for example degradation or losses of land and water resources and demand space for urban development [2]. Land use change in urban area influences by physical and human aspects that related with socio economic factors. Change detection for land use is important to determine the rapid changes happen in an area this have been widely used for planning and managing land [3]. Global environment changes can be identified with the identification of land use land cover elements [4]. Moreover, to have substantial development, state authorities need to have a tool to efficiently monitor land use land cover changes and planning towards future demand.

Therefore, decision makers, stakeholders and city citizen should understand the impact to the environment due to the changes that happen towards the land cover. This will ensure that urban green spaces

and its ecosystem is protected for better sustainable environment planning. Identification and extraction of information for change detection purpose can be done using digital classification.

Remote sensing has several advantages such as time saving, update map information, cover huge data with less time (more economic), digital image classification, variety dataset (multi resolution and multi temporal) and spectral information [1]. Conventional methods for land use land cover classification and mapping is a time consuming and high cost to process [5]. Mapping of land use and land cover changes are important for global environment monitoring and future development planning [6, 7]. Remote sensing technology has an advantage to produce accurate classified result of land use land cover over large area and combined with conventional field sampling data [8]. Furthermore, remote sensing is an advance tool that offers several classification approach for mapping of land use and land cover.

Remote sensing classification techniques has been used widely to determine the distribution of earth surface cover. Image classification had made great progress for land cover map production with the use of its advance classification algorithms [6]. The high resolution multispectral sensor contains a visual band and infrared layer that can represent the real sight of earth surface visualization, thus it can be used for classification of land information [3]. Most of the land use land cover classification were derived from remotely sensed data using pixel based method of classification mainly by supervised classification. Remote sensing classification algorithm that mostly used for land information mapping is the Maximum Likelihood classification (MLC) and Support Vector Machines (SVM) [9, 10]. In supervised classification approach, useful information categories are distinct first by identify the training sites and each pixel is classified based on statistical analysisthen their separability is examined accurately [11].

MLC is known to be the most popular supervised classification algorithm used widely in several application for field mapping [11]. Remote sensing dataset is widely used for land cover classification and monitoring land use changes using Maximum Likelihood classification approach, this classification approach has proven to produce high accuracies that more that 80% [12]. This classification using parametric approach which the spectral statistic data should be distributed normally [13, 14]. The MLC tool considers both the variances and covariance's of the class signatures when assigning each cell to one of the classes represented in the signature file [15].

Based from recent research in the classification field, Support Vector Machines (SVM) also known as supervised classification algorithm that has been used in the field of remote sensing with its primary advantage in generalization with limited training samples [7, 9, 16]. This classification approach has been used in several remote sensing datasets such as Landsat multispectral and Hyperspectral data [17]. SVM algorithm is a non-parametric that determines the boundaries in feature space by divide and differentiates the classes [18]. SVM has been proven a good classifier for complex man made infrastructures that have similar spectral signatures [19-21]. This classification algorithm able to provide higher accuracy in comparison to Maximum Likelihood classification for certain application [22-25]. This research work aims to access the classification accuracy of land use land cover using pixel based with MLC and SVM classification algorithm in urban area of Kuala Lumpur, by multi temporal Landsat imagery.

2. MATERIALS AND METHOD
2.1. Study area
The selected study area is located in Kuala Lumpur, Malaysia, this area is known as the well-developed area with high population density. The study area covers the coordinates of 3° 8' 51" N, 101° 41' 36" E (Figure 1). This area has experienced huge urbanization process due to economic and population growth and expected to enlarge it urban area in future. The total population of Kuala Lumpur in 2016 is 1.79 million people in an area of 243 km² [26].
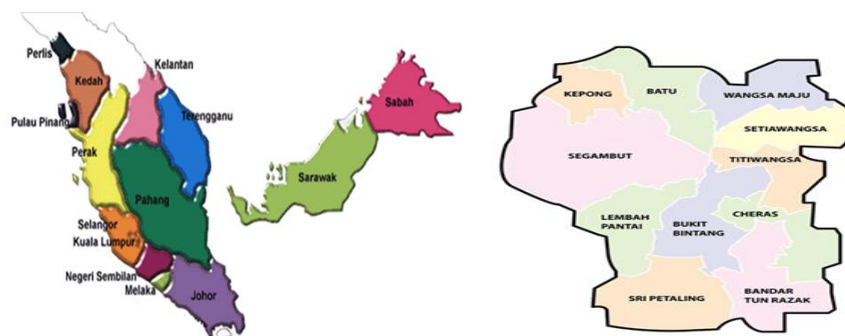

Figure 1. Study area, Kuala Lumpur (Source: Google)

2.2. Dataset

The data used for this research are mostly from secondary data such as satellite imageries: Landsat 5 TM (Thematic Mapper) for year 2001 and 2006, Landsat 7 ETM + (Enhanced Thematic Mapper plus) for year 2011 and Landsat 8 OLI (Operational Land Imager) for year 2016. The satellite imagery for this study were downloaded from the USGS Earth Explorer. All imageries were captured with different type of sensors with cloud cover less than 20% with 30m spatial resolution. This data is corrected for geometric and topographic errors. Others supporting is the current land use map of Kuala Lumpur provided by Town and Regional Planning department, this data is important for validation purpose.

2.3. Image Preprocessing

The main steps for image classification include selection of suitable classification method, selection of training samples, image pre-processing, feature extraction, selection of classification algorithm, post classification processing and accuracy assessment. Image classification is the stage of image analysis in which the multivariate quantitative measurement associated with each pixel is translated into a label from a pre-defined set (e.g., land use categories) [4]. The multi temporal images downloaded from the USGS free website is a standard product that have been geometrically and radiometrically corrected by the data providers. All the satellite images were geometrically corrected to the ortho-rectified level. For this study, the spatial detail, original pixel size and value for each image for specific bands are retain.

2.4. Classification of Images

As for the classification stage, a uniform supervised classification was applied on all images. Supervised classification was done by collecting training samples for land use land cover classes. Training areas was created accurately using polygon for each of the classes and classification was performed using Maximum Likelihood (ML) and Support Vector Machine (SVM) method. The classification was done by only considering five main classes: urban/built up area, vegetation, water body, low land and bare land. The main classes and their description are presented in Table 1. The overall image classification procedure was illustrated in Figure 2.
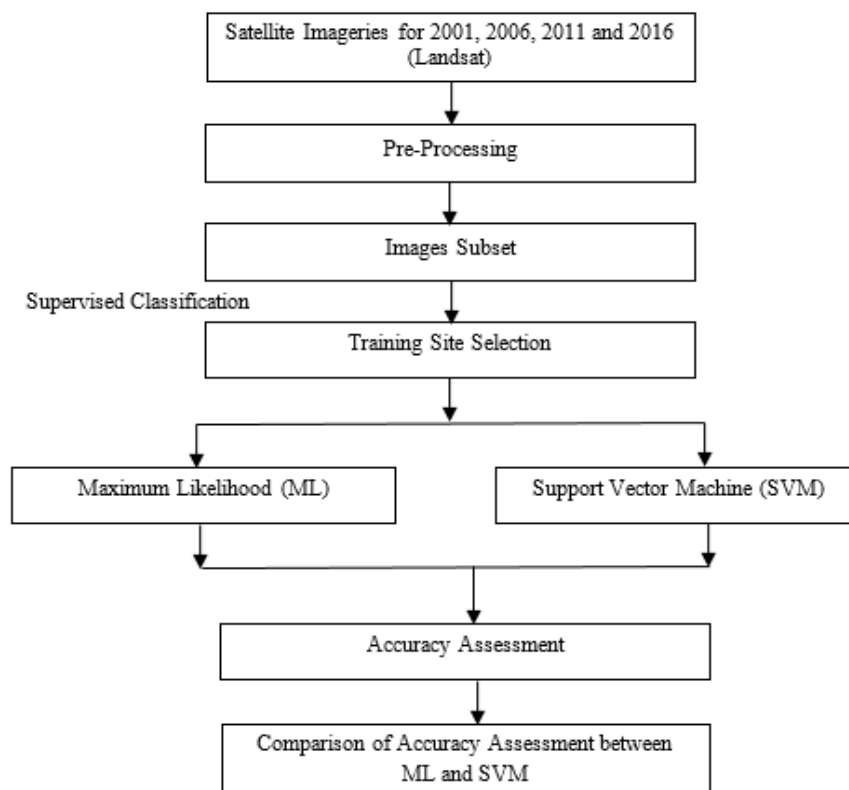

Figure 2. Flow chart of Image Processing

Table 1. Description of Land use Land Cover Classes

Classes	Description
Built-up area	Industrial, residential, commercial and services, transportation infrastructure, villages and settlement.
Vegetation	Trees, shrub lands and semi natural vegetation, deciduous, coniferous and mixed forest, gardens, inner city tree areas, park and playground, grass land and vegetable land.
Water Bodies	River, permanent open water, lakes, ponds, canals and reservoirs
Low land	Permanent and seasonal wetland, low-lying areas, marshy land, mudflats, all cultivated areas including urban agriculture, crop fields and rice paddies.
Fallow land	Sand in fillings, construction sites, developed land, excavation sites, solid waste landfills, open space, bare and exposed soils.

2.5. Accuracy Assessment

The final stage of the image classification process involves with an accuracy assessment process to evaluate the result from the classification stages [20]. This process is useful in evaluation of classification algorithm and able to determine the level of error that contribute by the image. The classification accuracy of each algorithm being used is expressed or shown in the form of confusion matrix. The accuracy measures for this paper tested the overall accuracy, user and producer accuracy and kappa coefficient.

3. RESULTS AND ANALYSIS

The final land use land cover maps for Kuala Lumpur in the year of 2001, 2006, 2011 and 2016 were presented in Figure 3 and Figure 4. The land use land cover maps show the major classes are built up area, vegetation, water body, low land and fallow land. The classification result was compare base on the corresponding accuracy assessment measures (overall, producer's and user's accuracies and kappa coefficient). The classification result was tested using confusion matrix. To perform the confusion matrix, training samples were needed. The five different time intervals (year) overall accuracy and kappa statistic produced using Maximum Likelihood were 2001 (83.46%, 0.73), 2006 (90.83, 0.85), 2011 (90.93, 0.855) and 2016 (89.38, 0.83) and using Support Vector Machine classifier were 2001 (91.65%, 0.86), 2006 (94.83%, 0.91), 2011 (93.43%, 0.89) and 2016 (92.83, 0.88). Therefore, both classification methods performed good results and can be used to classify land use and land cover accurately from multi temporal Landsat imagery. An evaluation of the accuracy of the classified images shows that the overall accuracies for support vector machine (SVM) classification is higher that Maximum Likelihood pixel-based classification results. This result indicates that support vector machine able to classify land use land cover in urban area accurately than maximum likelihood base from the accuracy results. This result supported by previous research that has evaluated various classification algorithm and concluded that SVM able to provide higher accuracy than MLC in land use land cover mapping [25].
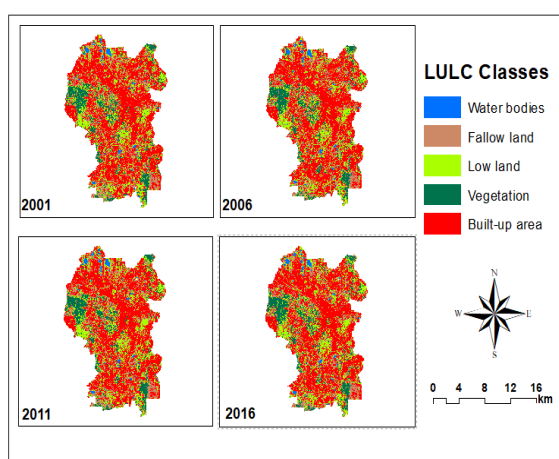

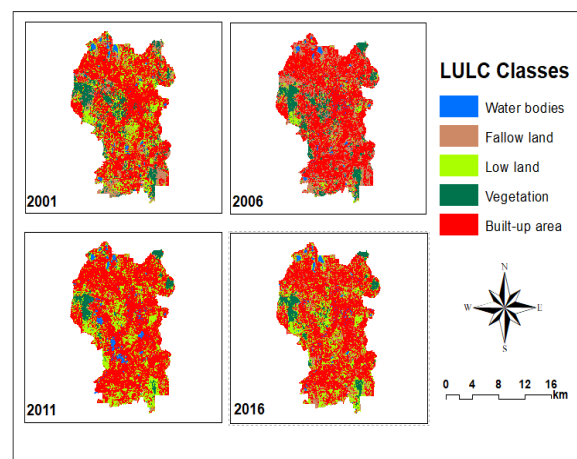Figure 3. Classified land use/land cover maps using Maximum Likelihood classifier for 2001-2016

Figure 4. Classified land use/land cover maps using Support Vector Machine classifier for 2001-2016

Table 2. Accuracy Assessment of Maximum Likelihood Classifier

Year	2001		2006		2011		2016	
Land Use/Land Cover	Producer's Accuracy (%)	User's Accuracy (%)	Producer's Accuracy (%)	User's Accuracy (%)	Producer's Accuracy (%)	User's Accuracy (%)	Producer's Accuracy (%)	User's Accuracy (%)
Urban/Built up Area	92.4	88.94	90.76	99.82	93.92	99.36	86.34	99.05
Vegetation	81.09	99.33	88.20	98.18	81.42	100	90.71	99.76
Low Land	88.66	54.24	86.55	64.98	94.96	51.95	92.86	72.22
Fallow Land	31.10	36.76	99.68	63.67	100	72.48	97.01	55.77
Water Bodies	94.79	94.79	98.96	97.94	97.92	96.91	98.96	98.96
Overall Accuracy	84.46%	90.83%	90.93%	89.38%				
Kappa Coefficient	0.73	0.86	0.85	0.83				

Table 3. Accuracy Assessment of Support Vector Machine Classifier

Year	2001		2006		2011		2016	
Land Use/Land Cover	Producer's Accuracy (%)	User's Accuracy (%)	Producer's Accuracy (%)	User's Accuracy (%)	Producer's Accuracy (%)	User's Accuracy (%)	Producer's Accuracy (%)	User's Accuracy (%)
Urban/Built up Area	94.47	99.77	95.96	99.31	97.23	99.77	92.98	98.53
Vegetation	85.46	97.63	97.60	97.07	83.39	99.87	91.69	99.53
Low Land	81.51	68.07	72.27	95.03	95.8	57.72	92.86	76.47
Fallow Land	99.00	65.63	96.17	70.99	100	86.98	93.41	69.03
Water Bodies	100	89.72	98.96	96.94	98.96	87.96	98.96	94.06
Overall Accuracy	91.65%	94.83%	93.43%	92.83%				
Kappa Coefficient	0.86	0.91	0.89	0.88				

4. CONCLUSION

As conclusion, remote sensing approach able to identify and classify urban land use land cover type and changes. Furthermore, remote sensing dataset and classification method can help in producing an accurate land use land cover map for several applications related with urban planning and development. The success of an image identification classification in remote sensing depends on several factors such as the availability of quality remotely sensed imagery, the framework of the classification procedure and the skills or experience by the analysts to execute the image processing procedure. The classification result will reflect in the accuracy assessment of the classification in the statistical table. In this study, the comparison between two different classification algorithms has been used to extract urban land use and land cover classes. The multi temporal remote sensing dataset of Kuala Lumpur, in 5 years' interval has been used (Year 2001, 2006, 2011 and 2016). From the results, it is proven that the level of accuracy from support vector machine method is higher than maximum likelihood method. Maximum likelihood probably unable to handle complex images therefore many of the pixels are not classified accordingly, as compared to support vector machine, this classifier able to overcome the mixed pixel problem and classified the pixel correctly. Selection of classification algorithm plays an important role to identify and classify every object in an urban area. Satellite remote sensing technology are useful for understanding the land use and land cover change dynamics. Therefore, this study would be time and cost-effective methods for urban planners and decision makers towards better planning for new development. According to this study the selection of classification algorithm plays is important to develop accurate land use land cover map for urban area and its surroundings. This information is useful for further monitoring and planning purposes of land use and land cover in urban area in future.

REFERENCES

[1] Ahmad, A. and Quegan, S., Analysis of maximum likelihood classification on multi spectral data. *Applied Mathematics and* Science, 6 pp. 6425-6436, 2012. [1]
[2] Aspinal, R. J. and Hill, M.J., Land Use Change, Science Policy and Management (New York: CRC Press Taylor and Francis Group), 2008.

[3]   Deilmai, B. R., Ahmad, B. &Zabihi, H., Comparison of two Classification methods (MLC and SVM) to extract land use land cover in Johor Malaysia. In IOP *Conference Series: Earth and Environment Science*. https://doi.org/10.1088/1755-1315/20/1/012052, 2014a.

[4]   Deilmai, B. R., Ahmad, B. &Zabihi, H., Comparison of two Classification methods (MLC and SVM) to extract land use land cover in Johor Malaysia. In IOP *Conference Series: Earth and Environment Science*. https://doi.org/10.1088/1755-1315/20/1/012052, 2014b.

[5]   Al-Ahmadi, F., Hames, A., Comparison of Four Classification Methods to Extract Land Use Land Cover from Raw Satellite Images for some Remote Arid area, Kingdom of Saudi Arabia. *Journal King AbdulAziz Univ. Sci.*, 20, 167-191, 2009.

[6]   BirohmatinAmalisana, R. and R.H., Land Cover Analysis bu Using Pixel-Based and Object-Based Image Classification Method in Bogor. *Earth and Environment Science*, 98(1), 1-7. https://doi.org/119.40.114.73, 2017.

[7]   Foody, G. M. and Mathur, A., A relative evaluation of multiclass image classification by support vector machines. *Geoscience and Remote Sensing*, IEEE Transactions on, 42, 1335-1343, 2004.

[8]   Lillesand, T.M. and Kiefer, R.W., Remote Sensing and Image Interpretation. 3rd Edition, John Wiley and Sons, Inc., Hoboken, 750, 1994.

[9]   Erener, A., Classification method, spectral diversity, band combination and accuracy assessment evaluation for urban feature detection. *International Journal of Applied Earth Observation Geoinformation*, 2013; 21, 397-408, 2013.

[10] Jensen, J.R.,Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd Edition, Prentice-Hall, Inc., New York, 316, 2004.

[11] Keuchel, J., Naumann, S., Heiler, M. and Siegmund, A., Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data. *Remote Sensing of Environment*, 86, pp. 530-541, 2003.

[12] Latif, Z.A., Zamri, I. & Omar, H., Determination of trees species using WorldView-2 data. IEEE 8th Int. Conf. on Signal Process Appl. and Technol. (ICSPAT), 383–387, 2012.

[13] Al-doski, J., Mansor, S., &MohdShafri, H. Z., Image Classification in Remote Sensing. *Journal of Environment and Earth Science*, *3*(10), 141–148, 2013.

[14] Sahebjalal, E. and K. Dashtekian, (2013). Analysis of land use-land covers changes using normalized difference vegetation index (NDVI) differencing and classification methods. *African Journal of Agricultural Research*, 8 (37): p. 4614-4622, 2006.

[15] Srivastava, P., Han, D., Miguel, A., Ramirez, R., Bray, M., and Islam, T.,: Selection of classification techniques for land use/land cover change investigation. *Adv Space Res.*50: 1250–1265, 2012.

[16] Kavzoglu, T. and Colkesen, I., A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, 11, 352-359, 2009.

[17] Fonteh, M., Theophile, F., Cornelius, M., Main, R., Ramoelo, A., & Cho, M., Assessing the Utility of Sentinel-1 C Band Synthetic Aperture Radar Imagery for Land Use Land Cover Classification in a Tropical Coastal Systems When Compared with Landsat 8. *Journal of Geographic Information System*, 8, 495-505. doi: 10.4236/jgis.2016.84041, 2016.

[18] Kavzoglu, T. and Colkesen, I., A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, 11, 352-359, 2009.

[19] Salleh, S.A., Latif, Z.A., Pradhan, B., Wan Mohd, W.M.N., and Chan, A., "*Functional relation of land surface albedo with climatological variables: a review on remote sensing techniques and recent research developments,*" Geocarto International 29 (2), 147-163 (2014).

[20] Zhu, Z., & Woodcock, C. E., Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, 144, 152–171, 2014.

[21] Otukei, J.R., &Blaschke, T., Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation*, 12, 527-31, 2010.

[22] Taati, A., Sarmadian, F., Mousavi, A., Pour, C.T.H. and Shahir, A.H.E., Land Use Classification Using Support Vector Machine and Maximum Likelihood Algorithms by Landsat 5 TM Images. *Walailak Journal of Science and Technology*, 12, 681-687, 2015.

[23] Markham, B. L., and Townsend, R. G., Land cover classification accuracy as a function of sensor spatial resolution *Proceedings of the 15th International Remote Sensing of the Environment* (Ann ArborUSA,) (University of Michigan Press) 1075-1090 11-15, 1981.

[24] Mondal, A., Kundu, S., Chandniha, S. K., Shukla, R. & Mishra, P., Comparison of Support Vector Machine and Maximum Likelihood Classification Technique using Satellite Imagery, 2012.

[25] Pal, M. and Mather, P., Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26, 1007-1011, 2005.

[26] M. M. S. MALAYSIA, "Department of Statistics, Malaysia," 2016.